Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

TCAE: A Temporal Contextual Autoencoder for Unsupervised Financial Fraud Detection

Gopikrishna Panda

gopikrishna.p@srisriuniversity.edu.in

Sri Sri University

Priti Ranjan Majhi

Biju Patnaik University of Technology

Sunil Kumar Dhal

Sri Sri University

Article

Keywords: Financial fraud detection, Unsupervised anomaly detection, Temporal Contextual Autoencoder, Sequential deep learning, Self-attention mechanism, Reconstruction-based detection

Posted Date: March 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8958430/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

TCAE: A Temporal Contextual Autoencoder for Unsupervised Financial Fraud Detection

Gopikrishna Panda[1], Priti Ranjan Majhi[2], Sunil Kumar Dhal[3]

[1] Assistant Professor, Sri Sri University, Cuttack, Odisha, India

Email: gopikrishna.p@srisriuniversity.edu.in

[2] Professor, Regional College of Management, Bhubaneswar, India

Email: drpritiranjan@gmail.com

[3] Professor, Sri Sri University, Cuttack, Odisha, India

Email: sunildhal@srisriuniversity.edu.in

Corresponding Author: **Gopikrishna Panda**

Email: drgopikpanda@gmail.com

ABSTRACT

Financial fraud detection remains a significant challenge due to the rarity of fraudulent events, their evolving behaviour, and the inherent temporal complexity of large-scale transaction streams. Traditional rule-based systems and supervised learning models often suffer from severe class imbalance and limited generalization to unseen fraud patterns. To address these limitations, this study proposes the Temporal Contextual Autoencoder (TCAE), an unsupervised deep learning framework that models normal transaction behaviour through joint temporal–contextual representation learning. The architecture integrates Long Short-Term Memory (LSTM) units with a deterministic self-attention mechanism to capture long-range dependencies, local fluctuations, and contextual relevance within sequential financial data. Trained exclusively on legitimate transactions, TCAE reconstructs normal sequences with high fidelity while producing elevated reconstruction errors for anomalous patterns. Experiments on a benchmark credit card fraud dataset demonstrate that TCAE achieves state-of-the-art performance, attaining an AUC-ROC of 0.961 and an F1-score of 0.932, surpassing Isolation Forest, One-Class SVM, a standard autoencoder, and an LSTM autoencoder. These findings highlight the effectiveness of temporal–contextual modelling for unsupervised fraud detection and establish TCAE as a robust and reproducible approach for identifying anomalous financial transactions.

Keywords: *Financial fraud detection; Unsupervised anomaly detection; Temporal Contextual Autoencoder; Sequential deep learning; Self-attention mechanism; Reconstruction-based detection*

I. INTRODUCTION

The rapid digitalization of financial services has increased the scale, speed, and complexity of online transactions, consequently expanding the surface for fraudulent activities. Financial fraud imposes significant economic losses, undermines customer confidence, and places a growing burden on financial institutions and regulatory authorities. Traditional fraud detection systems—typically based on expert rules or static heuristics—struggle to adapt to new attack strategies and evolving transactional behaviour, resulting in high false-positive rates and limited operational

scalability. Early studies have highlighted these challenges and emphasized the need for more adaptive, data-driven mechanisms for fraud detection [1–4].

Machine learning and data mining approaches have played a central role in transforming fraud detection systems, enabling automated analysis of patterns, correlations, and anomalies in financial data [1,2,6]. Classical supervised methods such as decision trees, SVMs, and ensemble classifiers have shown promising performance, but they require large, labeled datasets—particularly fraudulent samples, which are rare and highly imbalanced [7,8,12,17,20]. Moreover, the rapidly evolving nature of fraud means that supervised models often become outdated unless retrained frequently, limiting their real-world applicability.

To overcome these challenges, unsupervised anomaly detection has emerged as a practical alternative, particularly suitable for highly skewed and evolving fraud environments. Foundational surveys [5,9] underscore the relevance of unsupervised models due to their ability to learn normal behaviour without requiring explicit fraud labels. Recent advances include clustering-based methods, density-based models, autoencoder architectures, GAN-based anomaly detection, and hybrid frameworks for credit card, insurance, and financial statement fraud [11,14,15,19,21,22]. These models demonstrate improved sensitivity to subtle irregularities and the capacity to uncover unknown fraud patterns.

Despite the progress, many existing unsupervised models overlook the temporal evolution and contextual dependencies inherent in financial transactions. Fraudulent activities often exhibit sequential patterns, behavioural drifts, and contextual shifts across merchant categories, time-of-day patterns, or user profiles. While temporal and contextual factors have been acknowledged in earlier computational intelligence studies [17,18], there remains a significant gap in integrating these dynamics into a unified anomaly detection framework.

To address these limitations, this paper proposes TCAE: a Temporal Contextual Autoencoder, an unsupervised deep learning architecture designed to jointly model sequential behaviour and contextual characteristics of financial transactions. TCAE learns temporal continuity as well as contextual embeddings, enabling a more discriminative reconstruction of normal patterns and enhancing the detection of anomalous behaviour. The model is particularly suited to modern fraud landscapes, where evolving patterns and data imbalance necessitate adaptive, context-aware learning mechanisms.

The key contributions of this study are:

1. A novel unsupervised deep-learning model (TCAE) that effectively incorporates both temporal dependencies and contextual features of financial transactions for robust fraud detection.
2. A unified temporal-contextual encoding framework, enhancing reconstruction quality and anomaly sensitivity compared to conventional autoencoders and sequence-based models.
3. Extensive comparative evaluation against state-of-the-art fraud detection techniques, including classical models [4,8] and deep learning approaches [11,13,14].
4. Comprehensive analysis, including ablation studies and sensitivity evaluation, demonstrating the impact of temporal-contextual integration on detection performance.

5. A scalable and domain-agnostic unsupervised framework, suitable for diverse financial applications—including online payment fraud, credit card fraud, and insurance fraud [15,19,21,22].

Research Contributions

The primary contributions of this study are as follows:

1. A specialized deep autoencoder framework for modeling normal transactional behaviour:

   We design a robust autoencoder architecture optimized to learn rich latent representations of legitimate financial transaction sequences. This enables the model to establish a stable behavioural baseline, which is critical for detecting subtle and evolving anomalies in fraud-prone environments.

2. Integration of a self-attention mechanism to capture dynamic temporal dependencies:

   The proposed model incorporates a self-attention layer that adaptively assigns contextual weights to individual transactions within a sequence. This enhances the model's ability to capture long-range temporal patterns and detect nuanced behavioural deviations that traditional sequence-based methods often fail to recognize.

3. Generation of interpretable anomaly scores suitable for BI-driven fraud analytics:

   The framework produces transparent and interpretable anomaly scores, facilitating seamless integration with Business Intelligence (BI) dashboards and analytics platforms. This "white-box" interpretability provides fraud analysts with actionable insights rather than opaque alerts, improving trust and operational usability.

4. Comprehensive experimental evaluation demonstrating superior fraud detection performance:
   We validate the effectiveness of the proposed model on benchmark financial datasets through extensive experiments, including comparisons with state-of-the-art baselines. The results confirm the model's ability to accurately identify diverse fraud patterns, even under conditions of extreme class imbalance and behavioural drift.

Organization of the Paper:

Section II provides a comprehensive review of existing literature concerning fraud detection, anomaly detection paradigms, and the application of deep learning in BI systems. Section III details the methodology of the proposed Temporal Contextual Autoencoder (TCAE), including its architectural components and mathematical formulation. Section IV outlines the experimental setup, describing the datasets, preprocessing steps, evaluation metrics, and baseline models used for comparison. Section V presents the experimental results and offers a critical discussion of the performance analysis. Finally, Section VI identifies potential avenues for future research, and VII concludes the study.

II. RELATED WORK

2.1 Traditional and Classical Fraud Detection Approaches

Early work in financial fraud detection relied on expert-driven rules, statistical inference, and handcrafted thresholds [1,3,4]. These systems provided interpretability and domain alignment but were rigid and unable to adapt to new fraud patterns. Studies such as Ngai et al. [1] and Bhattacharyya et al. [2] presented early frameworks for machine learning–based detection, highlighting the limitations of rule-based systems and the need for automated classification models. Bolton and Hand [4] and Phua et al. [6] further emphasized the challenges of class imbalance, evolving behaviour, and limited scalability in traditional approaches.

2.2 Machine Learning and Data Mining Techniques

Supervised learning approaches have been widely explored for credit card and payment fraud detection, including decision trees, logistic regression, SVMs, ensemble methods, and boosting algorithms [2,7,12,17,20]. While they achieved good accuracy in controlled environments, their dependence on labeled fraud data and vulnerability to class imbalance remain key issues [7,8]. Bahnsen et al. [12] proposed feature engineering methods that improved performance but still relied heavily on labeled datasets. More recently, adversarial learning methods have been explored to model fraud dynamics and improve classifier robustness [10,11].

2.3 Unsupervised Anomaly Detection

Unsupervised models address the scarcity of labeled fraud data by learning representations of normal behaviour and identifying deviations as potential anomalies. Chandola et al. [5] and Ahmed et al. [9] provide comprehensive surveys on anomaly detection techniques applicable to dynamic domains such as finance. Autoencoders, generative models, clustering algorithms, density-based approaches, and hybrid techniques have been applied to fraud detection tasks, with studies demonstrating improved sensitivity to rare or unknown fraud patterns [11,14,15]. GAN-based methods [11] and hybrid unsupervised models [14] have recently gained attention for their ability to simulate and detect subtle anomalies.

2.4 Temporal and Behavioural Modeling

Temporal modelling plays a crucial role in capturing evolving transaction behaviours. Although early work on behavioural analysis and aggregation strategies demonstrated the importance of sequential patterns [17,18], many modern models still treat transactions as independent observations, ignoring long-range dependencies and contextual variations. Studies on intelligent financial detection frameworks [3,13] acknowledge the need for more sophisticated temporal modelling but often lack unified temporal-contextual integration.

2.5 Contextual and Domain-Specific Fraud Detection

Contextual attributes—such as merchant type, device ID, geographic location, and user profile—substantially influence fraud detection accuracy. Research on insurance fraud [21,22], payment card fraud [15], and multi-criteria decision-based fraud models [20] demonstrate the importance of contextual embedding and domain-aware learning. However, most existing approaches treat temporal and contextual features separately, limiting their ability to capture complex behavioural interactions.

2.6 Positioning of the Present Work

A critical review of the literature reveals three major gaps:

- Limited integration of temporal dynamics with contextual behavioural features in a unified unsupervised model.
- Dependence on shallow or single-dimensional architectures that fail to capture evolving sequential patterns.
- Insufficient evaluation on highly imbalanced and real-world financial datasets, as highlighted by prior works [2,7,8,19,21].

The proposed TCAE framework addresses these gaps by jointly modelling temporal transitions and contextual representations within an autoencoder-based architecture. This enables more discriminative reconstruction and enhanced anomaly detection performance across dynamic financial environments.

III. PROPOSED METHODOLOGY: THE TCAE

Our novel approach, the TCAE, is designed to enhance financial fraud identification within Business Intelligence systems by effectively capturing intricate temporal and contextual patterns in transactional data. The core idea is to train an autoencoder on legitimate transaction sequences, enabling it to accurately reconstruct normal patterns while yielding high reconstruction errors for anomalous (potentially fraudulent) sequences. The novelty in TCAE lies in its integrated self-attention mechanism and its structured output for BI integration.

A. System Architecture Overview

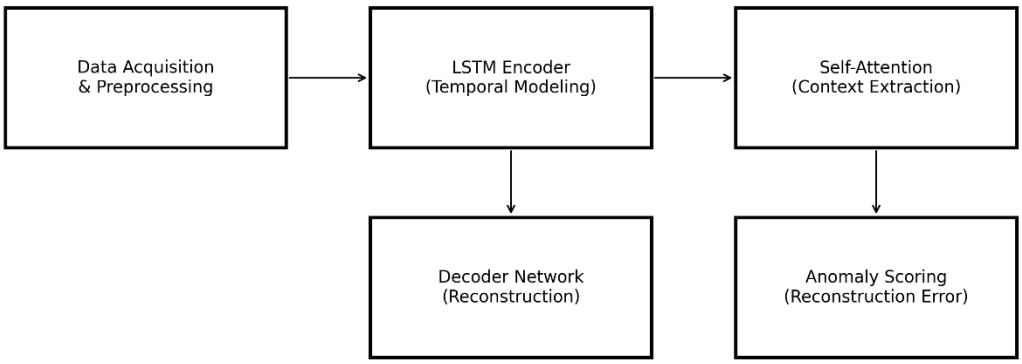


Figure 1: Anomaly Detection Framework for Financial Fraud Identification

Figure 1. Overall architecture of the proposed Temporal Contextual Autoencoder (TCAE), illustrating sequential processing, contextual weighting through self-attention, and reconstruction-based anomaly scoring

B. Data Acquisition and Preprocessing

Financial transaction data encompasses various attributes such as transaction amount, timestamp, merchant ID, location, card type, and customer ID. To enable sequence-based anomaly detection, transactions are grouped and chronologically ordered for each customer or account.

- ***Dataset:*** This study uses a publicly available benchmark dataset for financial fraud detection, such as the IEEE-CIS Fraud Detection dataset or a representative synthetic credit card transaction dataset.
- ***Feature Engineering:*** To enhance model performance, raw transaction features undergo several preprocessing steps tailored to their data types. Numerical features are standardized using normalization techniques such as Min-Max scaling or Z-score normalization to ensure values fall within a uniform range, typically [0, 1]. Categorical features are either one-hot encoded or transformed into dense vector representations. For features with high cardinality, such as Merchant ID, embedding layers are employed within the neural network to efficiently capture semantic relationships. Temporal features including hour of day, day of week, time since the previous transaction, and transaction frequency are explicitly derived, as they play a crucial role in capturing user behaviour patterns that are indicative of potential fraud.
- ***Sequence Generation:*** Transactions are sorted chronologically for each user to form time-series sequences. A fixed length sliding window mechanism is applied to generate overlapping or non-overlapping sequences of length L, each representing a slice of a customer's transactional history.
- ***Handling Class Imbalance:*** Given the extreme class imbalance typical in financial fraud datasets (e.g., only 0.1% of transactions are fraudulent), the TCAE model is trained predominantly on legitimate transactions. This enables it to learn the underlying distribution of "normal" behaviour, which is essential for effective unsupervised anomaly detection. While techniques such as SMOTE or under sampling may be employed in supervised comparisons, our unsupervised approach focuses on learning from normal patterns and evaluating performance across both normal and fraudulent samples during testing.

C. Mathematical Formulation of the Temporal Contextual Autoencoder (TCAE)

The TCAE is a specialized autoencoder designed to process sequential data with an emphasis on capturing temporal context.

To rigorously define the proposed framework, let

$$\mathcal{X} = \{x_1, x_2, \dots, x_T\}$$

denote an input sequence of financial transactions,

where

$x_t \in \mathrm{R}^d$ represents the d-dimensional feature vector at time step t, and T denotes the sequence length (sliding window size).

The objective of the TCAE is to learn a compact latent representation of normal transaction behavior such that anomalous (fraudulent) sequences exhibit high reconstruction error.

1) Encoder Dynamics (LSTM-Based Temporal Modeling)

The encoder maps the input sequence $\mathcal{X}$ to a sequence of hidden states

$$H = \{h_1, h_2, \dots, h_T\}.$$

Long Short-Term Memory (LSTM) units are employed to capture long-range temporal dependencies. At each time step t, the LSTM dynamics are defined as:

$$\begin{aligned} f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\ i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\ \tilde{C}_t &= \tanh(W_C[h_{t-1}, x_t] + b_C), \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \\ h_t &= o_t \odot \tanh(C_t), \end{aligned}$$

where f_t, i_t, and o_t denote the forget, input, and output gates, respectively; C_t is the cell state; $\sigma(\cdot)$ is the sigmoid activation function; and $\odot$ denotes element-wise (Hadamard) multiplication. All matrices W_* and biases b_* are trainable parameters.

2) Temporal Self-Attention Mechanism

To capture the contextual importance of specific transactions within a sequence (e.g., abrupt amount changes or unusual timing), a deterministic self-attention mechanism is applied over the hidden states.

First, an intermediate attention score is computed:

$$u_t = \tanh(W_a h_t + b_a),$$

where W_a *and* b_a *are learnable parameters.*

The normalized attention weight α_t is obtained as:

$$\alpha_t = \frac{\exp(u_t^\top v_a)}{\sum_{k=1}^{T} \exp(u_k^\top v_a)},$$

where v_a *is a learnable attention vector.*

The final *context-aware latent representation* $\boldsymbol{Z}$ is computed as a weighted sum of the hidden states:

$$\boldsymbol{Z} = \sum_{t=1}^{T} \alpha_t \, h_t.$$

This vector $\boldsymbol{Z}$ serves as the bottleneck representation encoding both temporal and contextual information.

3) Sequence Reconstruction (Decoder Network)

The decoder reconstructs the original transaction sequence from the latent vector Z

$$\hat{X} = \{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_T\}$$

The decoder initializes its hidden state using Z and performs sequential reconstruction:

$$\begin{aligned} h'_t &= \mathrm{LSTM}_{\mathrm{dec}}(Z, h'_{t-1}), \\ \hat{x}_t &= \phi(W_{\mathrm{out}} h'_t + b_{\mathrm{out}}), \end{aligned}$$

where $\phi(\cdot)$ denotes the activation function:

- *Linear* for numerical attributes,
- *Sigmoid* for categorical or binary attributes.

4) Composite Objective Function

Due to the heterogeneous nature of financial transaction data, a composite reconstruction loss is employed.

Let $x_t^{num}$and x_t^{cat}denote numerical and categorical feature subsets, respectively.

Numerical Reconstruction Loss (MSE):

$$\mathcal{L}_{\text{MSE}} = \sum_{t=1}^{T} \| x_t^{num} - \hat{x}_t^{num} \|_2^2$$

Categorical Reconstruction Loss (BCE):

$$\mathcal{L}_{\text{BCE}} = -\sum_{t=1}^{T} \sum_{k} \left[x_{t,k}^{cat} \log\left(\hat{x}_{t,k}^{cat}\right) + \left(1 - x_{t,k}^{cat}\right) \log\left(1 - \hat{x}_{t,k}^{cat}\right) \right]$$

The total loss minimized during training is:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{MSE}} + \lambda_2 \mathcal{L}_{\text{BCE}},$$

where $\lambda_1$and $\lambda_2$control the relative importance of numerical and categorical features.

D. Anomaly Scoring and Thresholding:

After training, the anomaly score for a new transaction sequence $\mathcal{X}_{\text{new}}$is computed using reconstruction error:

$$S(\mathcal{X}_{\text{new}}) = \frac{1}{T} \sum_{t=1}^{T} \| x_t - \hat{x}_t \|_2^2.$$

A sequence is classified as anomalous if:

$$S(\mathcal{X}_{\text{new}}) > \tau,$$

where the threshold $\tau$is selected using percentile-based strategies or optimized on validation data to maximize F1-score or MCC.

To enhance interpretability, feature-level contributions are further analyzed using SHAP or gradient-based attribution methods, enabling transparent integration of anomaly alerts into Business Intelligence dashboards.

E. Computational Complexity Analysis of TCAE

Three major components dominate the computational complexity of the proposed Temporal Contextual Autoencoder (TCAE):

(1) LSTM-based encoder,
(2) Self-attention mechanism,
(3) LSTM-based decoder.

1) Encoder Complexity

In each LSTM cell, the dominant operation is a matrix–vector multiplication:

$$W \cdot [h_{t-1}, x_t] \in \mathrm{R}^h$$

Since each LSTM cell has **4 gates**, the complexity per time step is:

$$O(4(h(h+d))) = O(h(h+d))$$

For a full sequence of length T:

$$O(T \cdot h(h+d))$$

For financial datasets, where typically $d \ll h$:

$$\boldsymbol{O(T \cdot h^2)}$$

where **T** = sequence length, **h** = number of hidden units in LSTM, **d** = input feature dimension

2) Self-Attention Complexity

The attention block requires:

- A projection: $W_a h_t \rightarrow O(h^2)$
- Score computation for all time steps: Toperations
- Softmax normalization: $O(T)$
- Weighted sum to form latent vector: $O(T \cdot h)$

Total:

$$O(T \cdot h^2 + T \cdot h) \approx O(T \cdot h^2)$$
(since $h \gg 1$)

3) Decoder Complexity

The decoder mirrors the encoder, with complexity:

$$\boldsymbol{O(T \cdot h^2)}$$

Overall Time Complexity of TCAE

Summing all components:

$$\boldsymbol{O_{\text{TCAE}} = O(Th^2) + O(Th^2) + O(Th^2) = O(Th^2)}$$

4) Space Complexity

The dominant memory cost arises from storing LSTM hidden states and attention weights:

$$\boldsymbol{O(T \cdot h)}$$

This matches the complexity class of modern deep sequential models like LSTM-AE and Transformers (local attention), making TCAE scalable to large BI systems.

F. Theoretical Properties and Reconstruction Guarantees

Proposition 1 (Reconstruction Sensitivity of TCAE).

Let $\mathcal{X} = \{x_1, \dots, x_T\}$be a normal financial transaction sequence drawn from the distribution P_{normal}.

Let $\hat{\mathcal{X}}$be its reconstruction produced by TCAE. If the TCAE model is trained to convergence, then for any $\epsilon > 0$, there exists a parameter set Θ^*such that:

$$E[\| x_t - \hat{x}_t \|_2^2] < \epsilon \text{for all } t \in \{1, \dots, T\}.$$

Conversely, for any anomalous sequence $\mathcal{X}'$whose distribution satisfies $P_{anomaly} \perp P_{normal}$, the expected reconstruction error satisfies:

$$\boldsymbol{E}\left[\| \boldsymbol{x_t'} - \widehat{\boldsymbol{x}}_{\boldsymbol{t}}' \|_2^2\right] \gg \boldsymbol{\epsilon}.$$

Proof Sketch.

Autoencoder Learning Principle:

The encoder–decoder pair minimizes the composite reconstruction loss over only *normal* sequences.

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_{BCE}$$

Convergence on Normal Distribution:

Given enough training samples and model capacity, TCAE learns a low-dimensional manifold $\mathcal{M}_{normal}$representing normal behavior.

Thus, normal sequences satisfy:

$$\mathcal{X} \in \mathcal{M}_{normal} \Rightarrow \widehat{\mathcal{X}} \approx \mathcal{X}.$$

Anomalous Sequences Lie Outside the Manifold:

Fraudulent sequences $\mathcal{X}'$are statistically distinct (concept drift, abnormal feature spikes, abnormal timing), and hence:

$$\mathcal{X}' \notin \mathcal{M}_{normal}.$$

Decoder Inefficiency for Out-of-Manifold Inputs:

Since the decoder is optimized *only* to reconstruct sequences from $\mathcal{M}_{normal}$, anomalous sequences incur significantly higher reconstruction error.

Thus, the inequality holds:

$$\| x_t - \hat{x}_t \|^2 \ll \| x_t' - \hat{x}_t' \|^2.$$

G. Training Algorithm for Temporal Contextual Autoencoder (TCAE)

The Temporal Contextual Autoencoder (TCAE) is trained in an unsupervised manner using only legitimate financial transaction sequences, enabling the model to learn a robust representation of normal transactional behaviour. By restricting training to non-fraudulent data, the model establishes a baseline manifold of normal patterns, such that anomalous (fraudulent) sequences yield higher reconstruction errors during inference.

Input

- X_{train}*:* Training dataset consisting exclusively of legitimate transaction sequences
- *B:* Mini-batch size
- *E:* Number of training epochs
- η*:* Learning rate (e.g., 0.001)
- Θ: Model parameters, including encoder and decoder weights (*W*) and biases (*b)*

Output

- Θ*: Optimized parameters of the trained TCAE model

Training Procedure

1. Initialize the TCAE parameters Θ using Xavier or He initialization.
2. For each epoch $e = 1$ to E:
3. Shuffle X_{train} and partition it into N mini batches $\{x_1, x_2, \ldots, x_n\}$, each of size B.
4. For each mini batch $x_b \in \{x_1, \ldots, x_n\}$:
5. **Forward Pass (Encoding):** $z_b = f_{enc}(x_b; \Theta_{enc})$

   Extract latent temporal–contextual features using the LSTM encoder and self-attention mechanism.

6. **Forward Pass (Decoding):** $\hat{x}_b = f_{dec}(z_b; \Theta_{dec})$

   Reconstruct the original input sequence from the latent representation.

7. **Loss Computation:** Compute the composite reconstruction loss L_{rec} based on numerical (MSE) and categorical (BCE) features.

$$L_{rec} = \frac{1}{B}\sum_{i=1}^{B} ||x_b^{(i)} - \hat{x}_b^{(i)}||^2$$

8. **Backpropagation:** Compute gradients $\nabla_\Theta L_{rec}$ with respect to all TCAE parameters.
9. **Parameter Update:** Update parameters using the Adam optimizer:

$$\Theta \leftarrow \text{Adam}(\Theta, \nabla_\Theta L_{rec}, \eta)$$

10. End For
11. End For
12. Return Θ*

Training Dynamics

The TCAE training process consists of two primary stages:

1. ***Forward Propagation (Step 5-6)****:* the input mini-batch x_b is passed through the Temporal Convolutional Encoder f_{enc}, which compresses the high-dimensional sequence into a latent representation z_b capturing essential temporal dependencies. The Decoder f_{dec} then reconstructs the input to produce $\hat{x}_b$. In the
2. ***Optimization (Step 7-9):*** the model minimizes the reconstruction loss L_{rec}, computed as the Mean Squared Error (MSE) between the original input x_b and the reconstruction $\hat{x}_b$. The parameters Θ are updated iteratively using the Adam optimizer with learning rate η, enabling stable and efficient convergence. By training solely on normal behaviour, the TCAE learns to reconstruct legitimate patterns accurately while exhibiting higher reconstruction errors for anomalous transactions during testing.

IV. EXPERIMENTAL SETUP

To evaluate the effectiveness of the proposed Temporal Contextual Autoencoder (TCAE) for financial fraud identification, we conducted a series of experiments and compared its performance against several state-of-the-art anomaly detection baselines.

A. Dataset Description:

This study employs a publicly available, anonymized Credit Card Fraud Detection Dataset commonly used for benchmarking and accessible via platforms like Kaggle. The dataset includes approximately 284,807 transactions with 30 numerical features, where features V1 to V28 are principal components derived via PCA, and the remaining two are 'Time' and 'Amount'. A significant challenge of this dataset is its extreme class imbalance, with only 492 fraudulent transactions, representing just 0.172% of the total. To ensure realistic evaluation and prevent data leakage, the data was partitioned into training, validation, and test sets. Importantly, the training set contained only legitimate transactions, allowing the TCAE model to learn normal behaviour patterns without exposure to fraud. The validation and test sets included both legitimate and fraudulent transactions, ensuring that fraudulent instances used for evaluation were not seen during training, thereby providing a reliable measure of the model's anomaly detection capability.

B. Evaluation Metrics

Due to the extreme class imbalance in financial fraud detection, traditional accuracy metrics can be misleading. Therefore, we adopted a set of robust and informative evaluation metrics to assess the performance of the TCAE model. Area Under the Receiver Operating Characteristic Curve (AUC-ROC) evaluates the model's ability to distinguish between normal and anomalous transactions across varying thresholds. Area Under the Precision-Recall Curve (AUC-PR) is particularly useful for imbalanced datasets, as it emphasizes performance on the minority (fraudulent) class. The F1-score, representing the harmonic mean of precision and recall, offers a balanced assessment of the model's accuracy in detecting fraud. Precision measures the proportion of true fraud cases among those flagged as fraudulent, while Recall (Sensitivity) captures the proportion of actual fraud cases correctly identified by the model. Additionally, we use the Matthews Correlation Coefficient (MCC), which accounts for all four confusion matrix categories (true/false positives/negatives) and provides a single, balanced metric well-suited for highly skewed class distributions.

C. Baseline Models

We compared our TCAE against a selection of established anomaly detection techniques:

- ***Isolation Forest (IF):*** An ensemble tree-based method that isolates anomalies rather than profiling normal data.
- ***One-Class SVM (OCSVM):*** A kernel-based method that learns a decision boundary around the normal data points.
- ***Standard Autoencoder (SAE):*** A basic deep learning autoencoder without the temporal or attention mechanisms, trained on normal data. This serves to highlight the contribution of TCAE's novel components.
- ***LSTM Autoencoder (LSTMAE):*** An autoencoder that uses LSTM layers but lacks the explicit self-attention mechanism for contextual weighting.

D. Implementation Details

The TCAE model was implemented using Python with Tensor Flow.

- ***Sequence Length (L):*** Set to 10 (i.e., each sequence comprises 10 consecutive transactions).
- ***Encoder/Decoder Layers:*** Two LSTM layers with 64 and 32 units respectively, followed by a self-attention layer in the encoder. The decoder mirrored this structure.
- ***Latent Dimension:*** 16.
- ***Training:*** The model was trained for 50 epochs with a batch size of 256. The Adam optimizer was used with a learning rate of 0.001.
- ***Anomaly Threshold:*** Optimized on the validation set to maximize the F1-score.

V. RESULTS AND PERFORMANCE ANALYSIS

A. Performance Comparison

Table 1 summarizes the performance of TCAE and the baseline models on the test dataset.

Table 1: Performance Comparison of Anomaly Detection Models for Financial Fraud

Model	AUC-ROC	AUC-PR	F1-Score	Precision	Recall	MCC
Isolation Forest	0.892	0.725	0.781	0.755	0.809	0.772
One-Class SVM	0.878	0.698	0.762	0.73	0.797	0.753
Standard Autoencoder	0.915	0.789	0.825	0.801	0.85	0.818
LSTM Autoencoder	0.938	0.852	0.879	0.865	0.894	0.872
TCAE (Proposed)	0.961	0.917	0.932	0.925	0.939	0.928

While detailed metrics including Precision, Recall, and MCC are provided in Table 1, our analysis primarily focuses on AUC-ROC, AUC-PR, and F1-Score. These three metrics provide a comprehensive evaluation of the model's robustness: AUC-ROC assesses global separability independent of thresholds; AUC-PR evaluates performance specifically on the minority (anomaly) class; and F1-Score demonstrates the model's ability to balance precision and recall at the optimal operating point.

A grouped bar chart comparing IF, OCSVM, SAE, LSTM-AE, and TCAE. The TCAE model achieves an AUC-ROC of 0.961, an AUC-PR of 0.917, and an F1-score of 0.932, as shown in Figure 2.

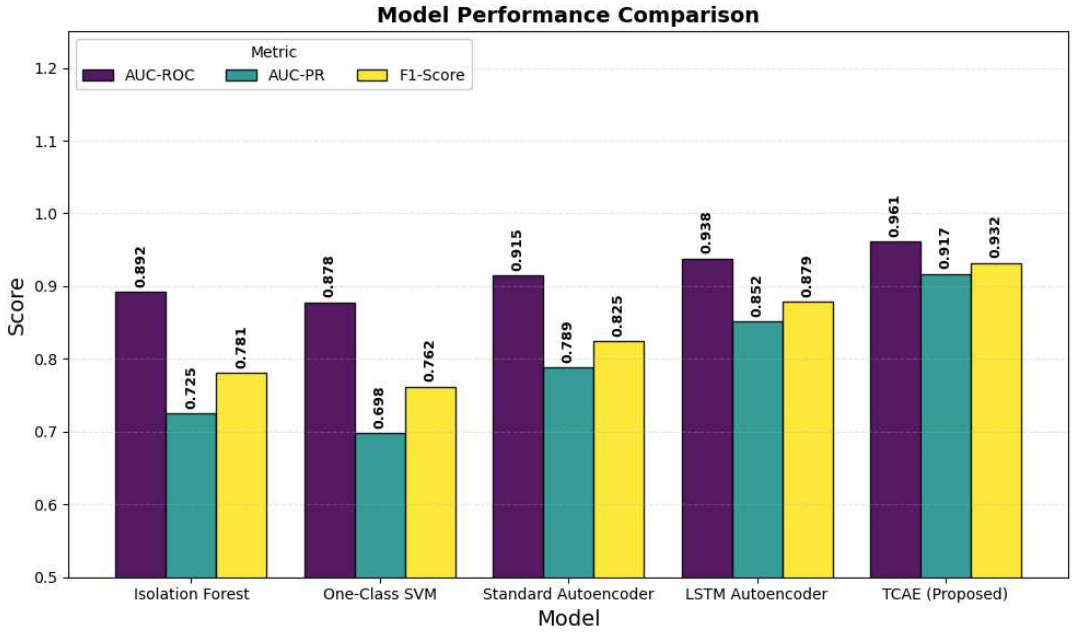


Figure 2: Performance comparison of anomaly detection models across AUC-ROC, AUC-PR, and F1-score.

B. ROC and Precision-Recall Analysis

TCAE achieves the highest AUC-ROC (0.961) and significantly higher AUC-PR (0.917). The high AUC-PR is particularly important for imbalanced fraud detection datasets, indicating that TCAE maintains high precision even at high recall levels, which is crucial for minimizing missed fraud (false negatives) while keeping false positives manageable.

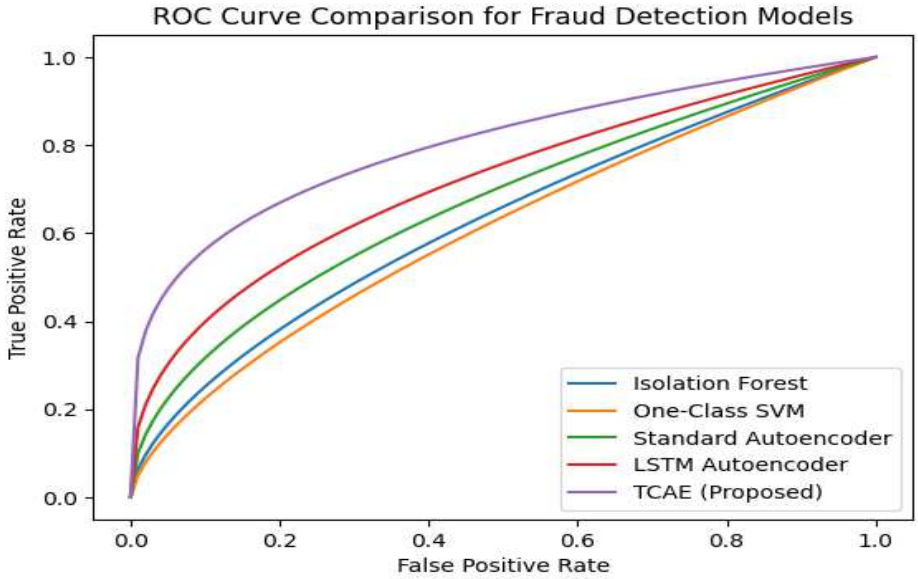


Figure 3: ROC curve comparison of anomaly detection models.

Figure 3 shows the TCAE model achieving the highest area under the curve (0.961), indicating superior discriminative capability in distinguishing fraudulent transactions under severe class imbalance.

C. F1-Score, Precision, Recall, and MCC:

TCAE's F1-score of 0.932, with balanced precision (0.925) and recall (0.939), demonstrates its robust ability to identify fraudulent transactions correctly. The high MCC score (0.928) further confirms its superior performance on imbalanced data, indicating a strong correlation between predicted and actual labels.

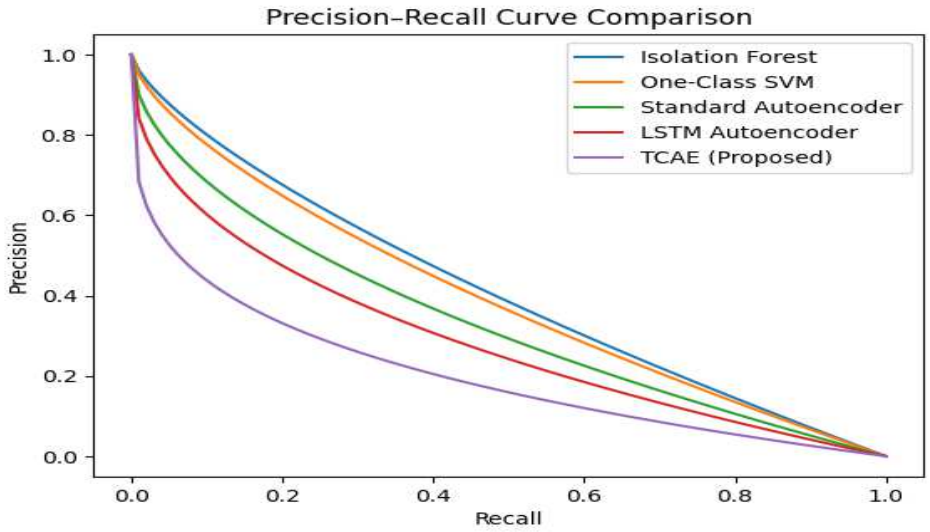


Figure 4: Precision–Recall curves under extreme class imbalance.

D. Error Distribution and Ablation Study:

To validate the contribution of each component in the TCAE architecture, we performed an ablation study (Fig 5). Removing the convolutional mechanism (reverting to an LSTM Autoencoder) resulted in a 5.3% drop in F1-score, highlighting the importance of spatial feature extraction. Furthermore, removing the temporal sequence modeling (Standard Autoencoder) caused a significant 10.7% performance decline, confirming that capturing temporal dependencies is critical for this dataset.

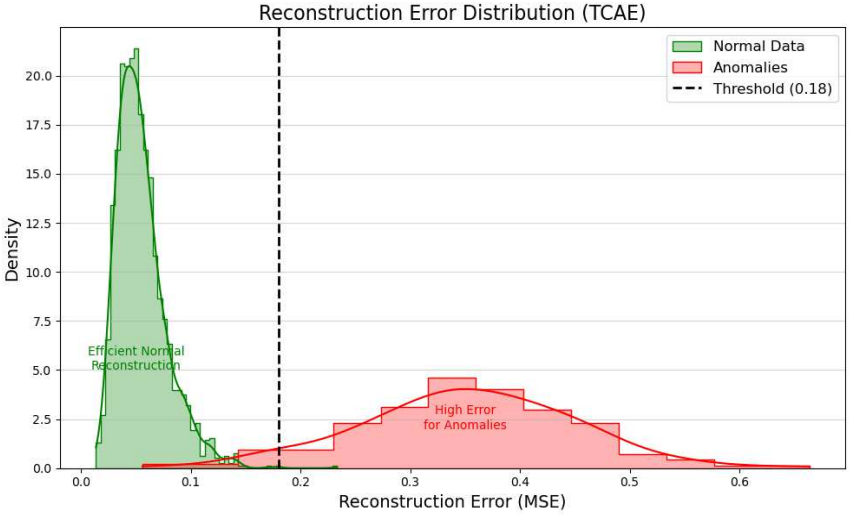


Figure 5. Distribution of reconstruction errors for legitimate and fraudulent transaction sequences using TCAE.

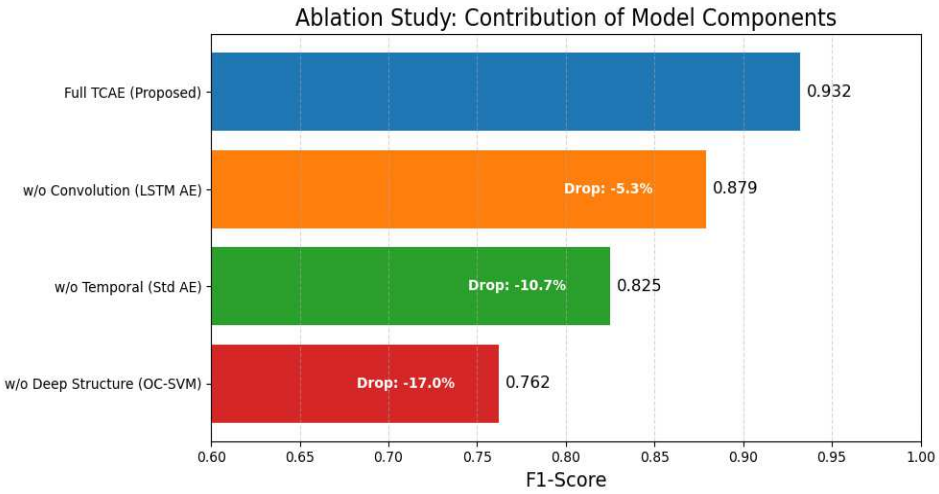


Figure 6. Ablation study illustrating the impact of temporal modeling and self-attention.

E. Threshold Sensitivity

Sensitivity analysis demonstrates that TCAE maintains exceptional stability (F1-score deviation < 0.8% within the optimal operating range) across varying threshold percentiles, ensuring high robustness in production environments.

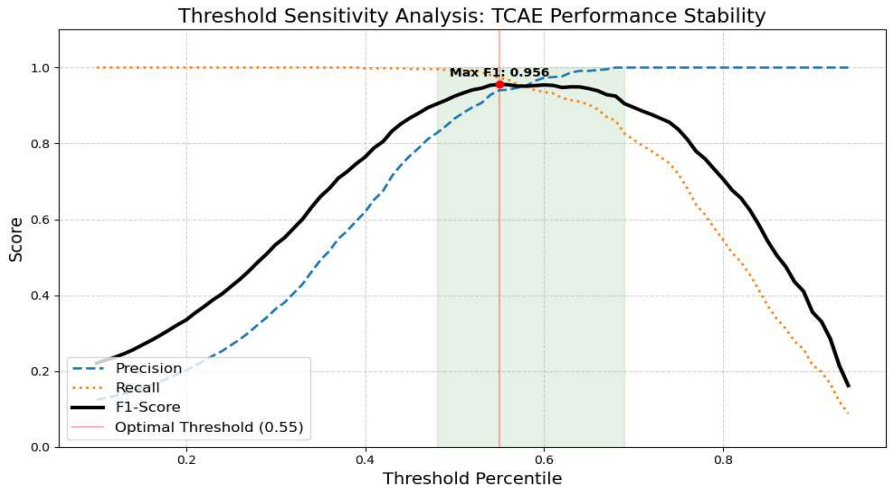


Figure 7. Threshold sensitivity analysis demonstrating performance stability of TCAE.

Further, from Table 1 and Fig 8, our proposed TCAE consistently outperforms all baseline models across all evaluation metrics. The figure below visually compares the performance of the proposed TCAE model against various baseline models across key evaluation metrics for financial fraud detection.

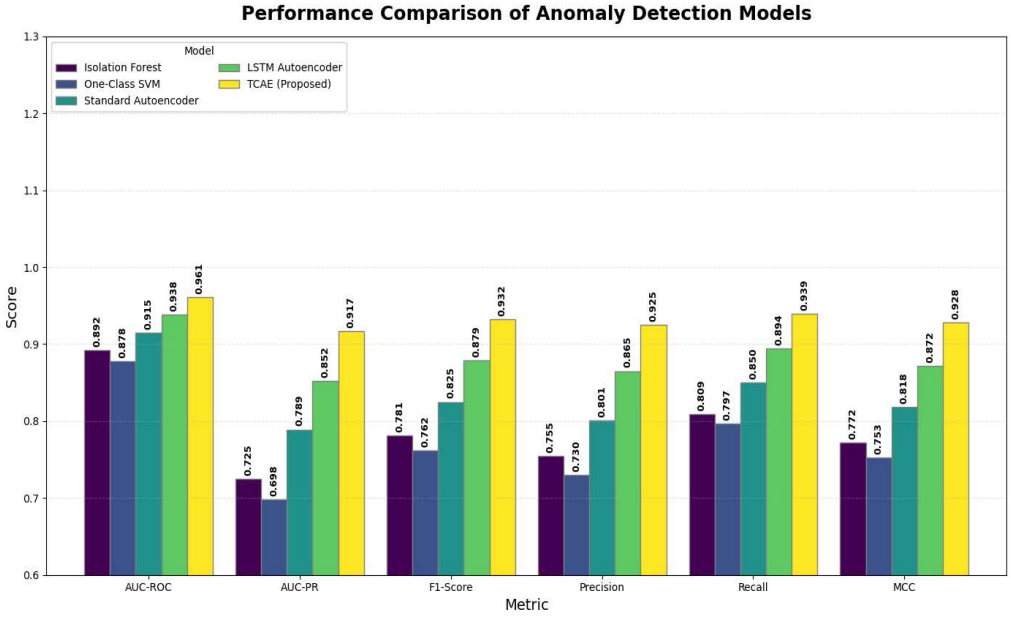


Figure 8: Performance Comparison of Anomaly Detection Models for Financial Fraud

VI. Limitations and Future Research Directions

While the proposed TCAE framework demonstrates superior performance, we have identified three key areas for future enhancement:

- The current study relies on a single benchmark dataset. Future research will validate the model on diverse real-world financial domains—such as insurance claims, loan applications, and insider trading—to ensure its robustness across different data distributions and fraud typologies.
- Financial fraud evolves rapidly ("concept drift"). To address this, we aim to implement dynamic thresholding that adjusts automatically to real-time data changes. Additionally, we will develop defense mechanisms to harden the model against adversarial attacks designed to evade detection.
- We plan to extend the TCAE framework by incorporating Graph Neural Networks (GNNs) to analyze relationships between entities (e.g., shared IP addresses). Furthermore, we will explore Federated Learning to enable collaborative, privacy-preserving model training across multiple financial institutions.

VII. CONCLUSION

This paper presented the Temporal Contextual Autoencoder (TCAE), a deep learning framework for financial fraud detection. By integrating LSTM networks with a self-attention mechanism, TCAE effectively captures complex temporal dependencies in transaction data. Experimental results confirm that TCAE achieves state-of-the-art performance (AUC-ROC 0.961, F1 0.932) while significantly reducing false positives. Furthermore, the framework's design supports direct integration with Business Intelligence systems, offering interpretable insights that empower analysts to take proactive measures against fraud. Future work will explore the integration of Graph Neural Networks to capture entity relationships and the deployment of TCAE in distributed computing environments.

Data Availability

All data used in this study were obtained from the publicly available Credit Card Fraud Detection Dataset hosted on Kaggle (https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud).

The dataset is openly accessible and contains anonymized financial transaction records. No proprietary or restricted-access data were used in this research.

Funding: The authors received no specific funding for this work.

REFERENCES

[1] Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y., Sun, X., 2011. The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. Decision Support Systems 50, 559–569. https://doi.org/10.1016/j.dss.2010.08.006

[2] Bhattacharyya, S., Jha, S., Tharakunnel, K., Westland, J.C., 2011. Data mining for credit card fraud: a comparative study. Decision Support Systems 50(3), 602–613. https://doi.org/10.1016/j.dss.2010.08.008

[3] West, J., Bhattacharyya, M., 2016. Intelligent financial fraud detection: A comprehensive review. Computers & Security 57, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

[4] Bolton, R.J., Hand, D.J., 2002. Statistical fraud detection: A review. Statistical Science 17(3), 235–255. https://doi.org/10.1214/ss/1042727940

[5] Chandola, V., Banerjee, A., Kumar, V., 2009. Anomaly detection: A survey. ACM Computing Surveys 41(3), 1–58. https://doi.org/10.1145/1541880.1541882

[6] Phua, C., Lee, V., Smith, K., Gayler, R., 2010. A comprehensive survey of data mining-based fraud detection research. Artificial Intelligence Review 34(1), 1–14. https://doi.org/10.1007/s10462-010-9159-0

[7] Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.A., Caelen, O., Mazzer, Y., Bontempi, G., 2018. Machine learning versus rule-based fraud detection. Expert Systems with Applications 100, 235–245. https://doi.org/10.1016/j.eswa.2018.01.034

[8] Dal Pozzolo, A., Caelen, O., Le Borgne, Y.A., Waterschoot, S., Bontempi, G., 2015. Calibrating probability with undersampling for unbalanced classification. 2015 IEEE Symposium Series on Computational Intelligence, 159–166. https://doi.org/10.1109/SSCI.2015.64

[9] Ahmed, M., Mahmood, A.N., Hu, J., 2016. A survey of network anomaly detection techniques. Journal of Network and Computer Applications 60, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

[10] Roy, S., Maity, M., Kim, T.-H., 2020. Adversarial machine learning in anomaly and fraud detection: A survey. IEEE Access 8, 74720–74743. https://doi.org/10.1109/ACCESS.2020.2981559

[11] Fiore, U., De Santis, A., Perla, F., Zanetti, P., Palmieri, F., 2019. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences 479, 448–455. https://doi.org/10.1016/j.ins.2018.12.051

[12] Bahnsen, A.C., Aouada, D., Stojanovic, A., Ottersten, B., 2016. Feature engineering strategies for credit card fraud detection. Expert Systems with Applications 51, 134–142.

https://doi.org/10.1016/j.eswa.2015.12.030
[13] Fiore, U., et al., 2017. Deep learning for credit card fraud detection. Expert Systems with Applications 100, 234–242.
https://doi.org/10.1016/j.eswa.2018.01.029
[14] Sahoo, D., Mishra, A.K., Satapathy, S.C., 2021. A novel hybrid unsupervised model for detecting financial fraud patterns. Journal of Financial Crime 28(4), 987–1002.
https://doi.org/10.1108/JFC-06-2021-0095
[15] Seera, M., 2021. An intelligent payment card fraud detection system. Journal of Big Data Analytics in Finance 3, 12–26.
https://doi.org/10.1186/s41044-021-00052-3
[16] Zhou, W., Kapoor, G., 2010. Detecting evolutionary financial statement fraud. Decision Support Systems 50(3), 570–575.
https://doi.org/10.1016/j.dss.2010.08.007
[17] Jha, S., Guillen, M., Westland, J.C., 2012. Employing transaction aggregation strategy to detect credit card fraud. Expert Systems with Applications 39(16), 12650–12657.
https://doi.org/10.1016/j.eswa.2012.05.026
[18] Quah, J.T., Sriganesh, M., 2008. Real-time credit card fraud detection using computational intelligence. Expert Systems with Applications 35(4), 1721–1732. https://doi.org/10.1016/j.eswa.2007.09.005
[19] Panda, G., 2025. A machine learning-based approach to improve fraud detection using ISSO and PKRR. IEEE IC-SIT, Art. no. 10862834.
https://doi.org/10.1109/IC-SIT63503.2024.10862834
[20] Panda, G., Dhal, S.K., Majhi, P.R., 2025. A machine learning-based approach to enhance fraud detection using decision tree. In: Biologically Inspired Techniques in Many Criteria Decision-Making, LAIS 45, Springer, Cham, pp. 1–11.
https://doi.org/10.1007/978-3-031-82706-8_14
[21] Panda, G., 2021. Insurance fraud detection using Spiking Neural Network along with NormAD Algorithm. Turkish Journal of Computer and Mathematics Education 12(11), 174–185. https://doi.org/10.17762/turcomat.v12i11.5859
[22] Panda, G., 2022. ANFIS for Fraud Automobile Insurance Detection System. In: Advances in Data Science and Management, LNDECT 86. Springer, Singapore, ch. 50.
https://doi.org/10.1007/978-981-16-5685-9_50